SIO 210 CSP: Data analysis methods
L. Talley, Fall 2016

1. Sampling and error
2. Basic statistical concepts
3. Time series analysis
4. Mapping
5. Filtering
6. Space-time data
7. Water mass analysis

- Reading: DPO Chapter 6
  - Look only at the less mathematical parts
  - (skip sections on pdfs, least squares, EOFs and OMP except to note that the material is there)
  - 6.1, 6.2, 6.3.1, 6.4,
  - 6.6.2, 6.7.1, 6.7.2

1. **Sampling** and error: definitions

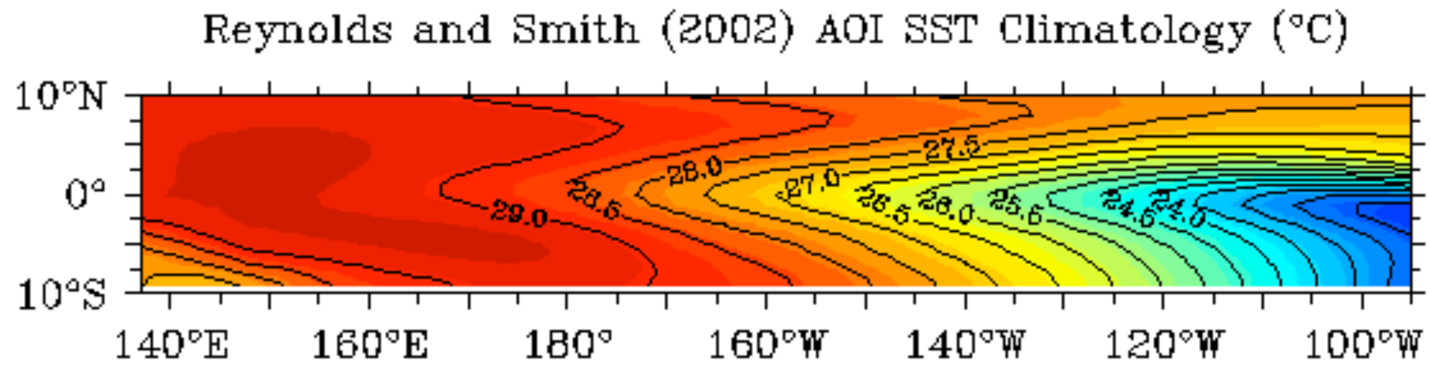
- Sampling (DPO Section 6.1)
  - Synoptic sampling
  - Climatology

  - Mean
  - Anomaly (difference between synoptic measurement and a mean

1. **Sampling**: mean and anomaly

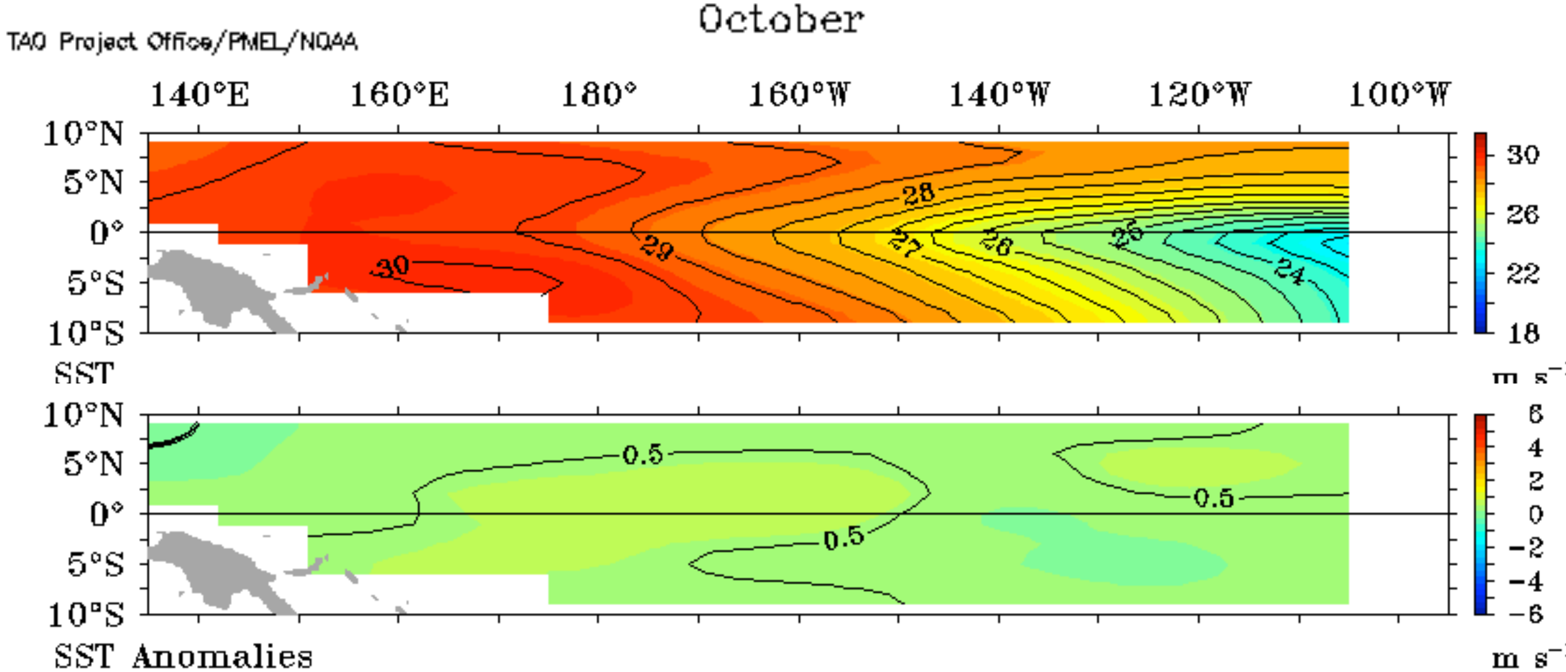
Climatology (mean): mean field based on large data base, usually covering many years

Climatology (Reynolds October Mean)



Observed Field (Oct. 2012)

Anomaly: Oct. 2012 Minus Reynolds October climatology



http://www.pmel.noaa.gov/tao/

1. Sampling and **error**

DPO Section 6.2

(PRECISION)

Random error**:** fluctuations (in either direction) of measured values due to precision limitations of the measurement device. Random error is quantified by the variance or standard deviation.

(ACCURACY)

Systematic error (bias): offset, high or low, which cannot be determined through statistical methods used on the measurements themselves.

An oceanographic example:

Two or more technical groups measure the same parameters (e.g. temperature, nutrients or oxygen, etc). The mean values the groups obtain differ because of differences in methods, chemical standards, etc. Error can only be evaluated by comparison of the two sets of measurements with each other or with an absolute standard.

2. Basic statistical concepts: Mean, variance, standard deviation (DPO Section 6.3.1)

Mean: $$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Anomaly: $x = x_{mean} + x_{anomaly}$; therefore $x_{anomaly} = x - x_{mean}$

Variance: $$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N} (x_i-\bar{x})^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N} (x_i)^2 - \frac{1}{N}\left(\sum_{i=1}^{N} x_i\right)^2\right]$$

Standard deviation of the measurements: σ (characteristic of the phenomenon)

Standard error: $\sigma/\sqrt{N}$ (characteristic of the SAMPLING since it depends on N)

3. Time series analysis (DPO 6.3.1, 6.3.3)

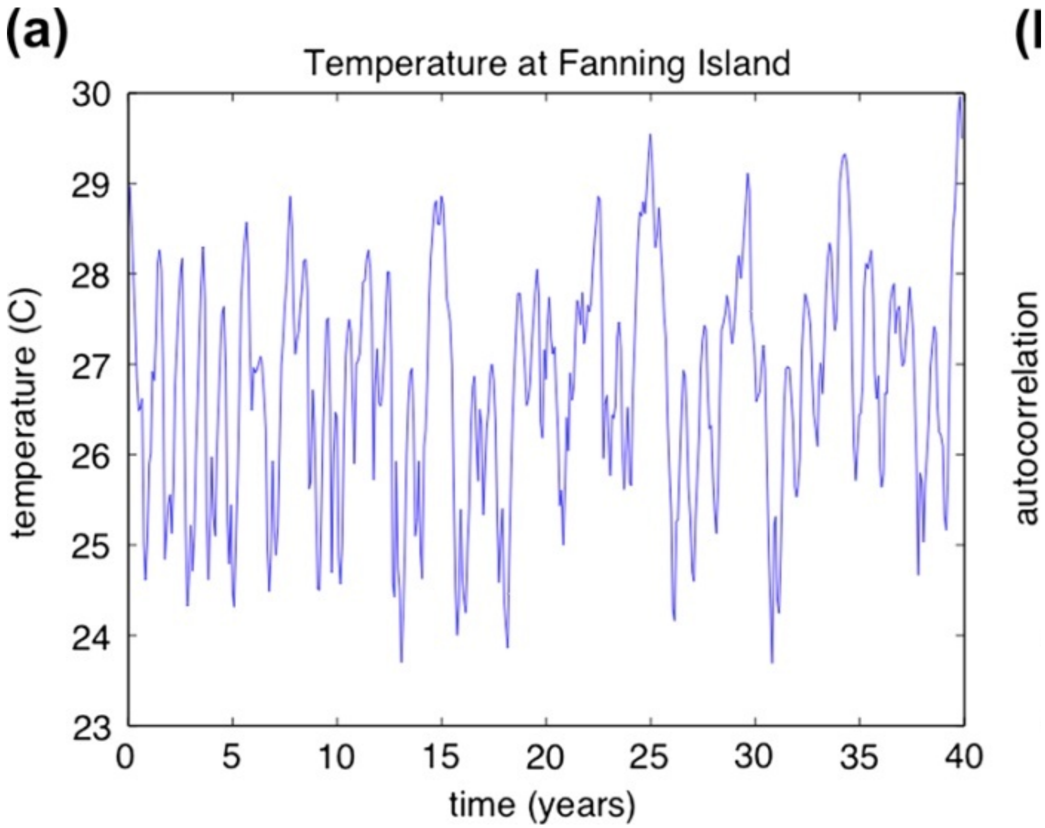
A time series is a data set collected as a function of time

Examples: current meter records, sea level records, temperature at the end of the SIO pier

Some common analysis methods:

A. Display the data (simply plot)! (always useful)
B. Mean, variance, standard deviation, etc. (example not shown)
C. Covariance and correlation of different time series with each other
D. Spectral (Fourier) analysis

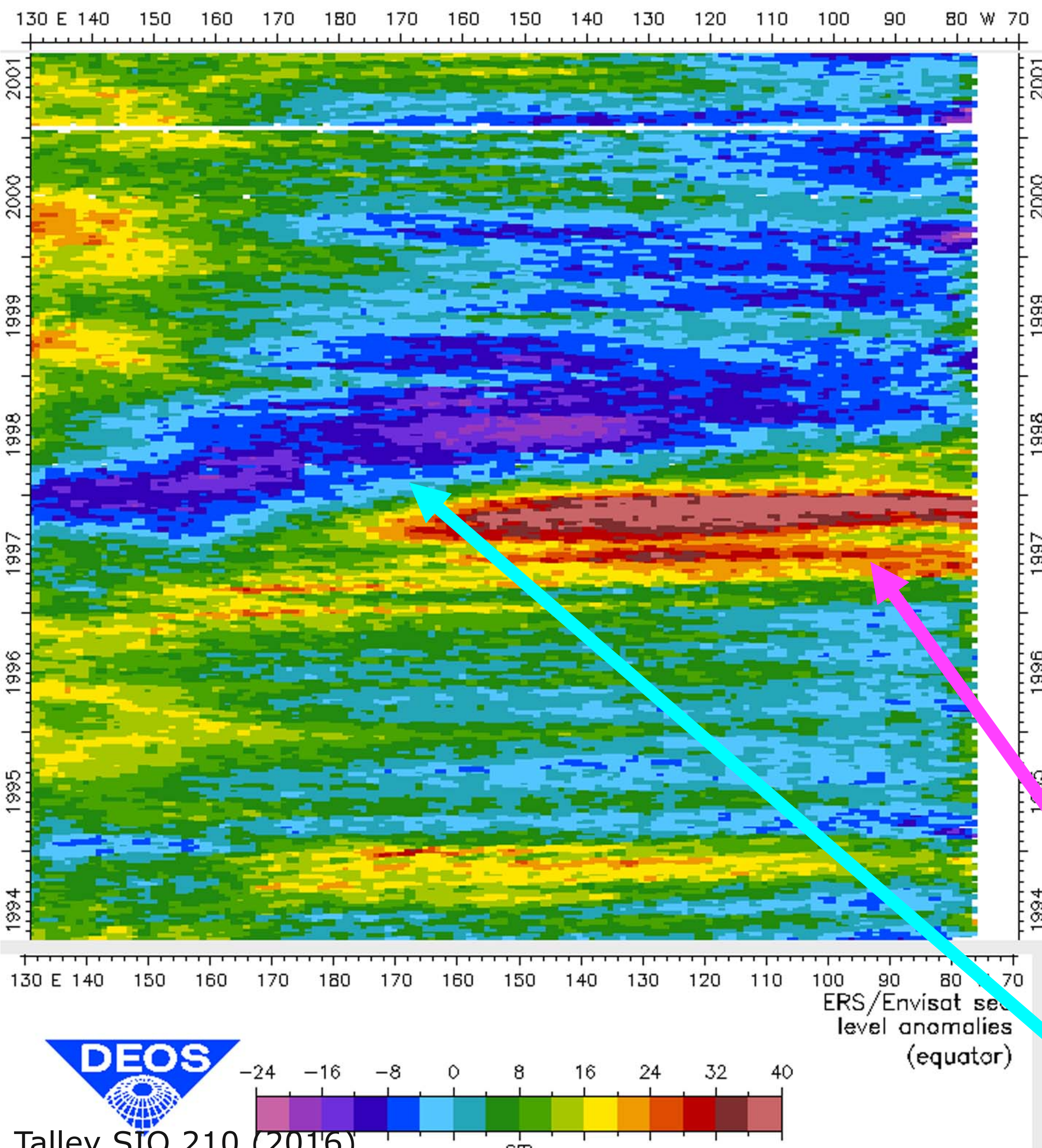
3.A. Display time series



(a) Time series of temperature at Fanning Island (Pacific Ocean) from the NCAR Community Ocean Model.

DPO FIGURE 6.2

3.A. Display time series: Hovmöller diagram (time on one axis, space on the other)



Equatorial Pacific sea level height anomaly from satellite

(High sea level anomaly - El Nino!)

(Low propagating west to east)

3.C. Time series analysis: covariance (DPO section 6.3.3)

Covariance and correlation: integral of the product of two time series, can be with a time lag. (Autocorrelation is the time series with itself with a time lag.)

- **Integral time scale**: time scale of the autocorrelation (definitions vary) – crude definition might be at what time lag does the autocorrelation drop to zero (which is the decorrelation timescale) (First calculate the autocorrelation for a large number of time lags.)

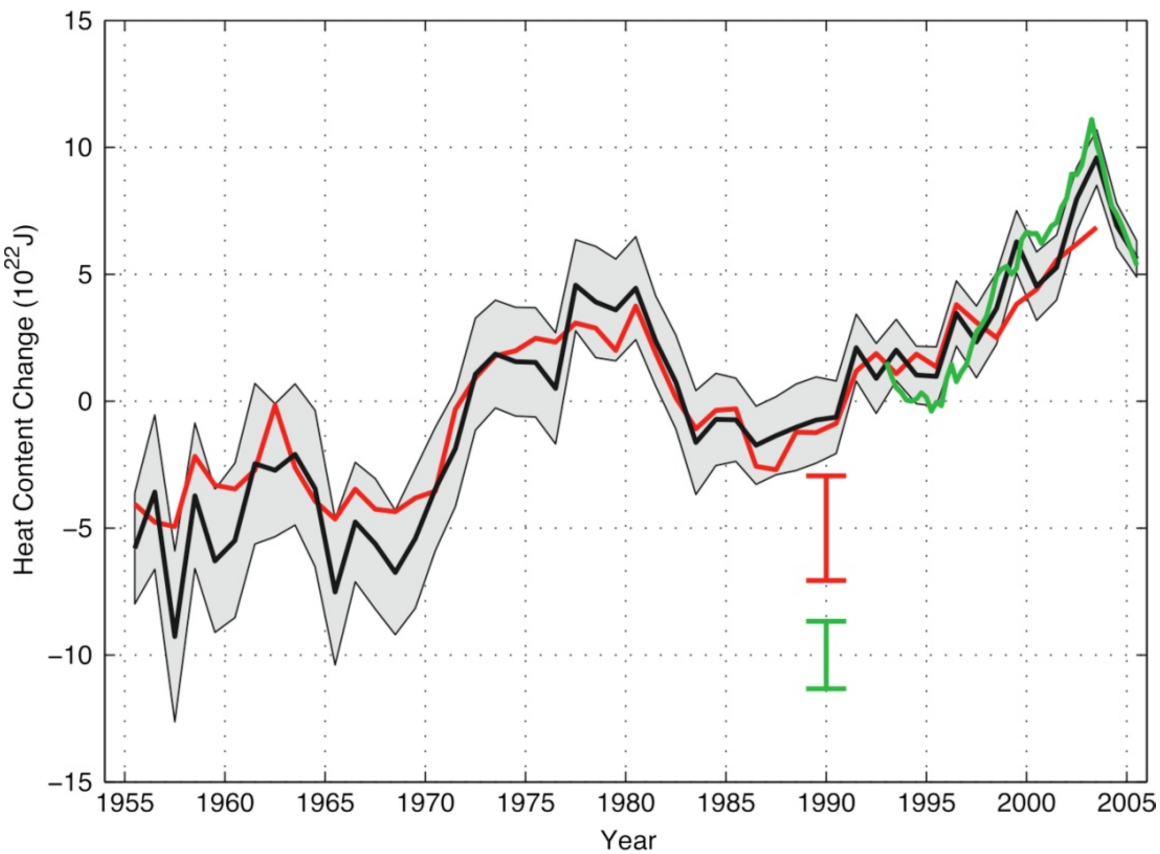
- **Degrees of freedom**: total length of record divided by the integral time scale – that gives how many realizations of the phenomenon you have. Good to have at least 10.

3.C. Time series analysis: confidence intervals (DPO Section 6.3.3)

- **Confidence intervals**: based on degrees of freedom and assumptions about statistical distribution

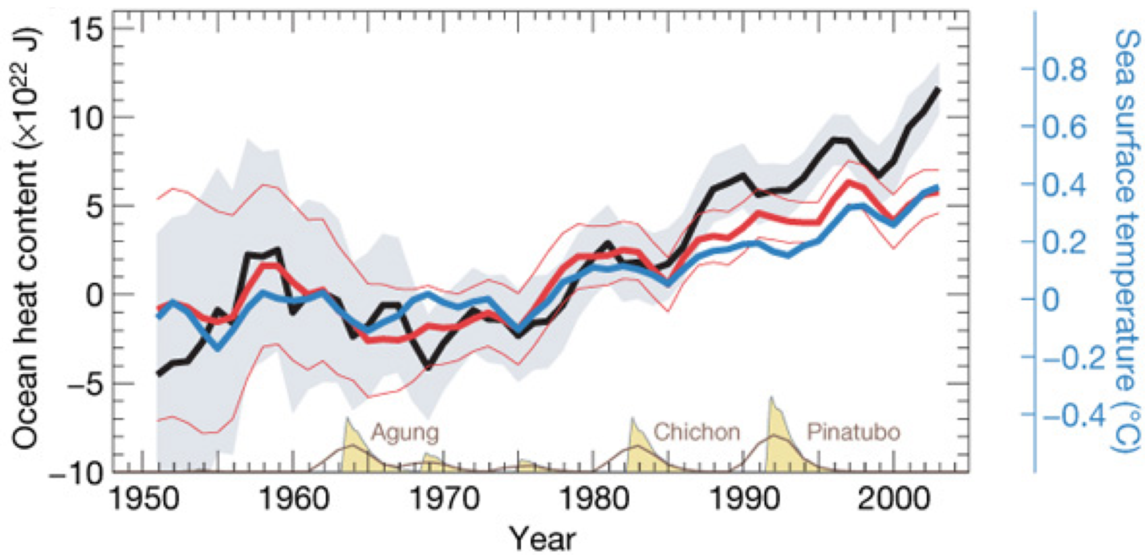
- **95% confidence interval**: 95% probability that a value is within the standard error of the mean (times a t-test quantity that you look up).

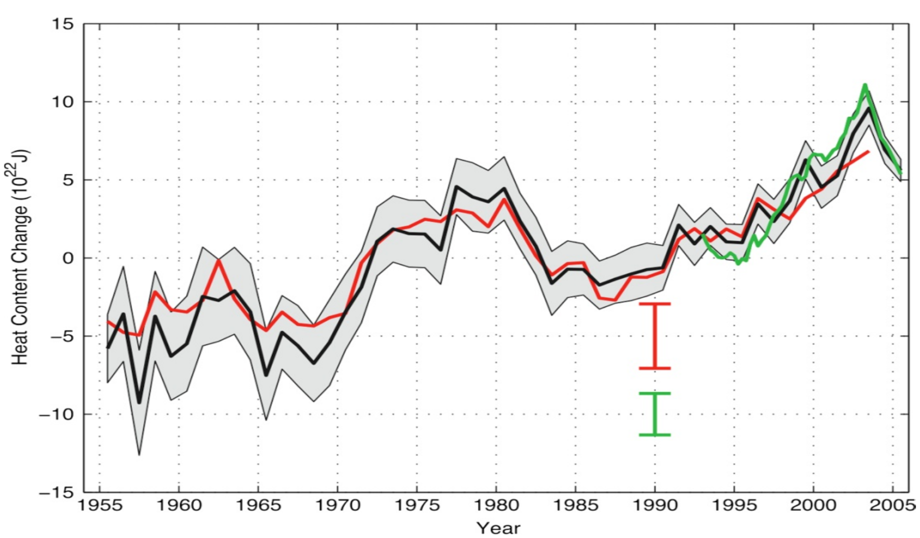
3.C. Time series analysis: confidence intervals (DPO Section 6.3.3)



Example of time series with confidence intervals. Global ocean heat content (10^{22} J) for the 0 to 700m layer, based on Levitus et al. (2005a; black curve), Ishii et al. (2006; full record gray curve and larger error bar), and Willis et al. (2004; darker gray after 1993 and shorter error bar). Shading and error bars denote the 90% confidence interval. Compare with Figure S15.15 seen on the textbookWeb site from Domingues et al. (2008) which uses improved observations

3.C. Time series analysis: confidence intervals (DPO Section 6.3.3)





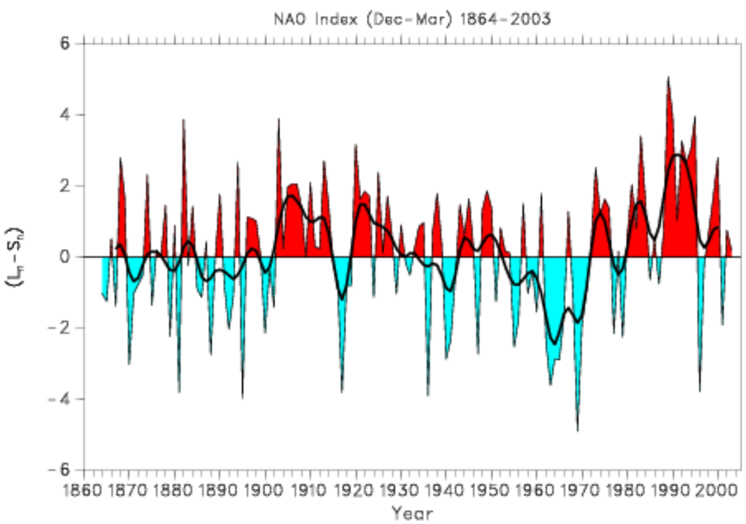
Talley SIO 210 (2016)

Figure S15.17 on the textbook Website from Domingues et al. (2008) which uses improved observations

Bottom line: uncertainty estimates are only as good as the underlying data!

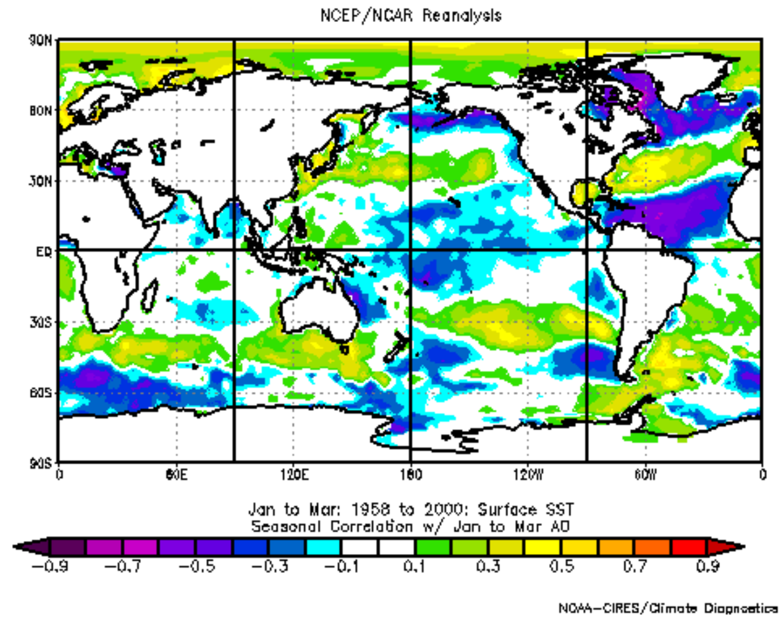
3.C. Time series analysis: correlation example

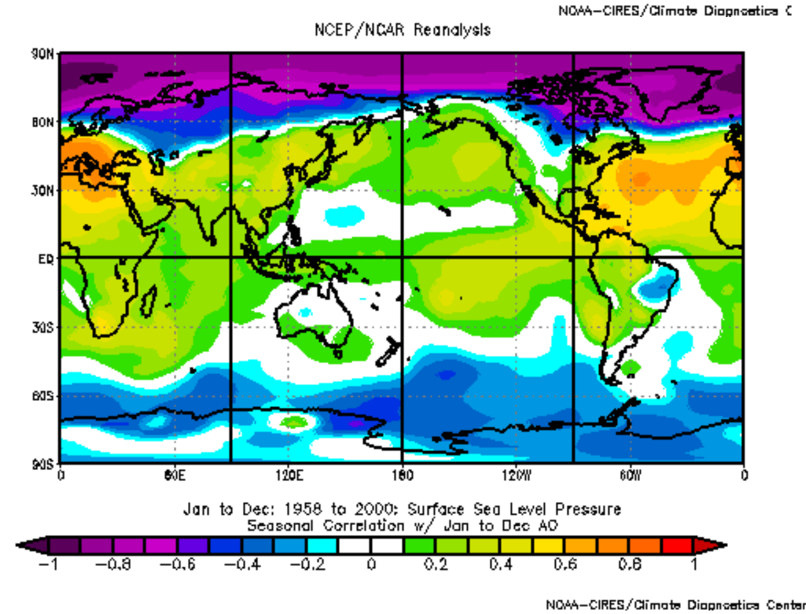
Example of correlation of different time series with each other



North Atlantic Oscillation index (time series)

Correlation with surface temperature and with surface pressure (done at each point, so each lat/lon gives a time series to correlate with NAO index)





Similar to DPO S15.2 and others S15 figures

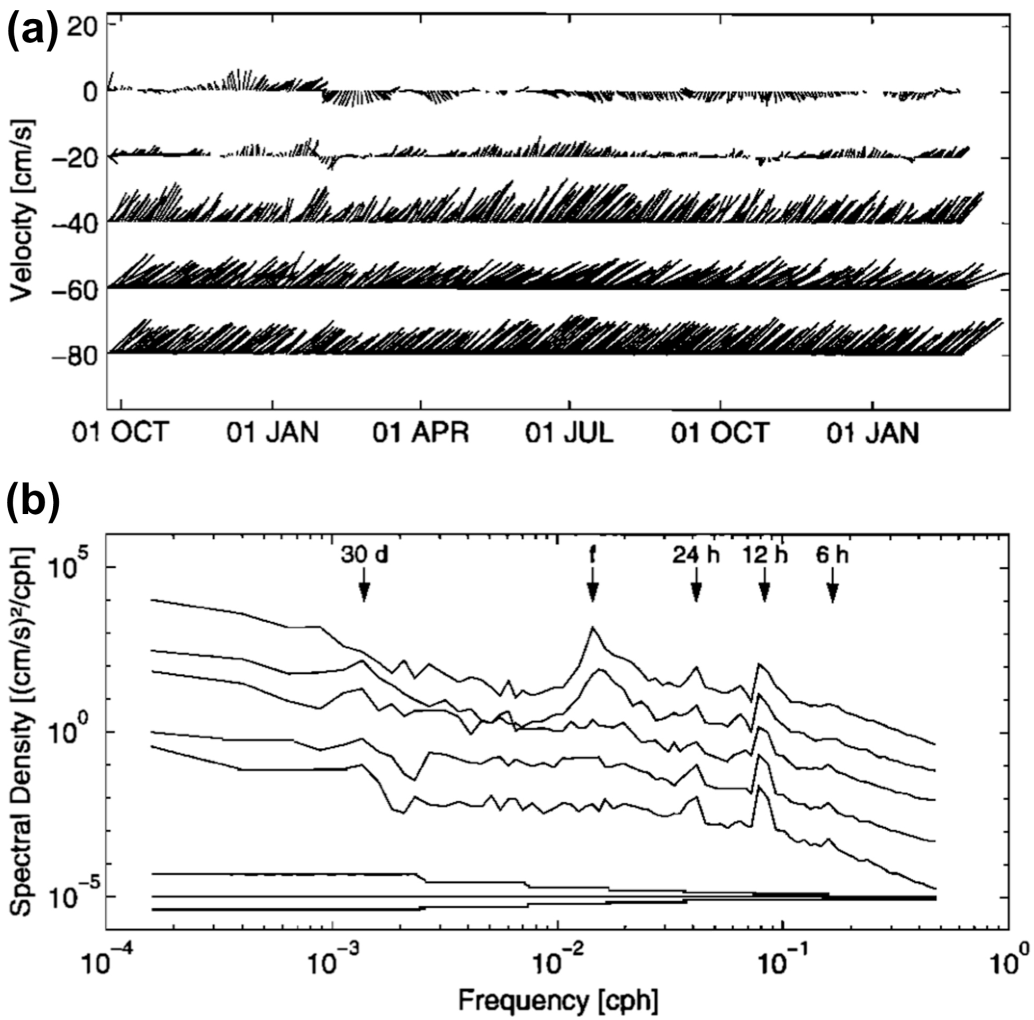
3.D. Time series analysis: spectral analysis (DPO Section 6.5.3)

Spectral (Fourier) analysis of a time series is used to determine its frequency distributions.

Underlying concept: any time series can be decomposed into a continuous set of sinusoidal functions of varying frequency.

Spectrum: amplitude of each of the frequency components. Very useful for detecting things like tides, inertial motions, that have specific forcing frequencies.

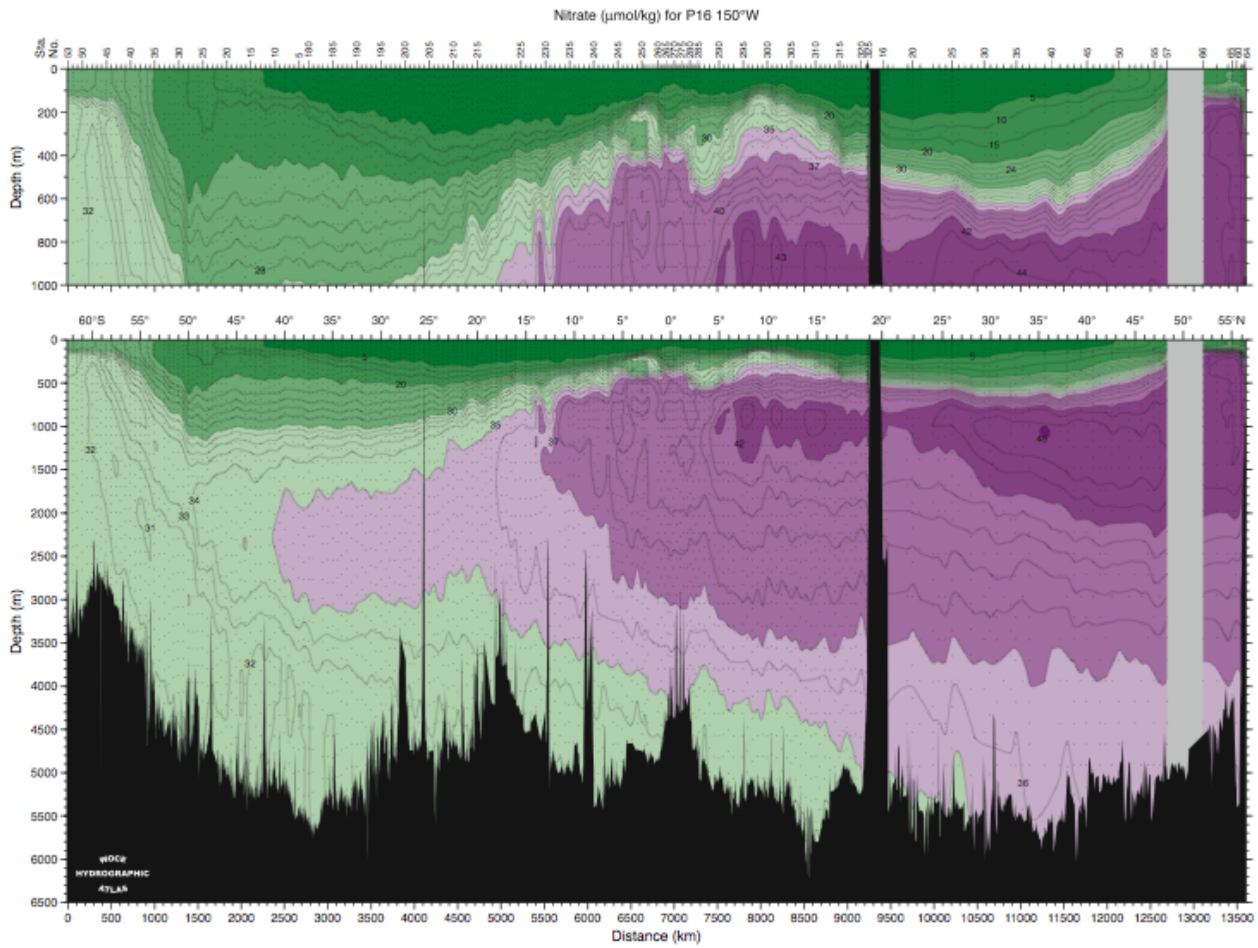
3.D. Time series analysis: spectrum
(DPO Section 6.5.3)



Example of time series, spectra, and spectral confidence intervals. (a) Velocity (cm/sec) stick plot, lowpassed at 100 hours, from 5 deep current meters at different depths on one mooring in the Deep Western Boundary Current in Samoan Passage (see Figure 10.16). The vertical direction is along the passage axis. (b) Spectra from the same current meters, offset by one decade. The 95% confidence intervals are shown at the bottom. *Source: From Rudnick (1997).*

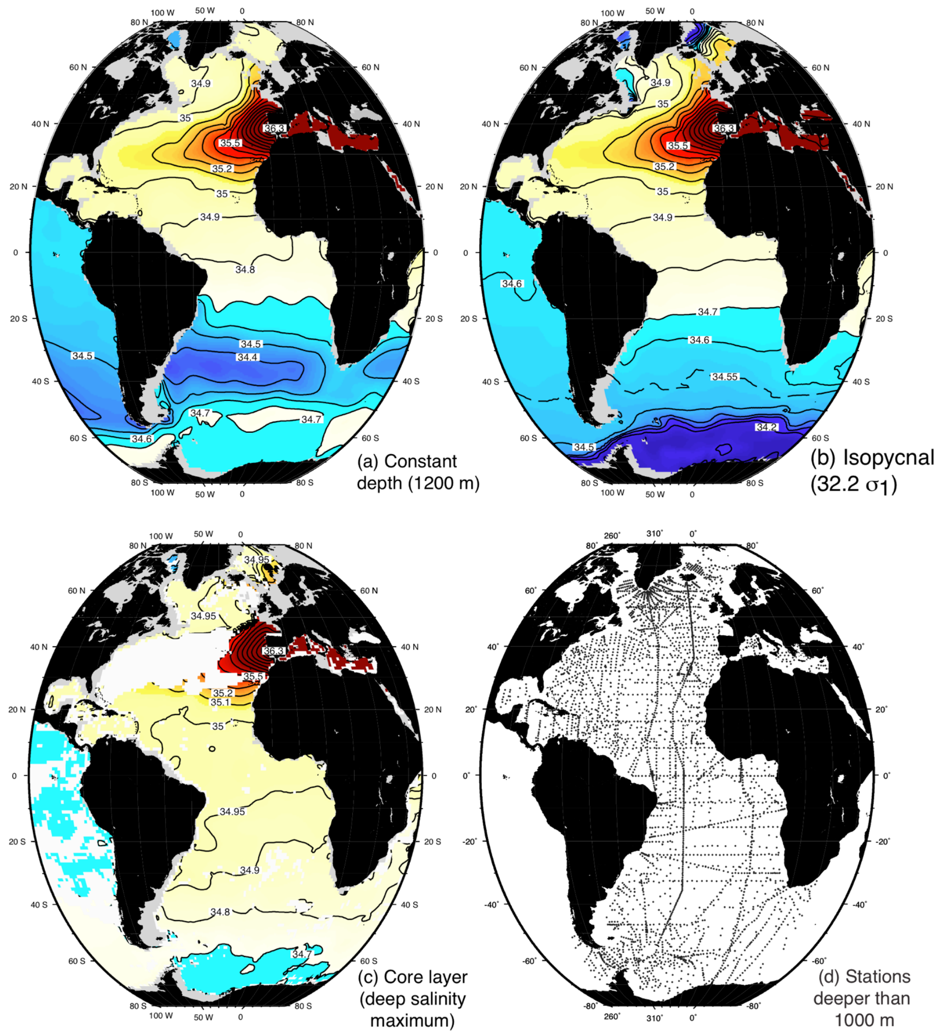
4. Oceanographic sampling: objective mapping in the vertical (DPO Section 6.4.1)

Discrete bottle samples: objectively mapped to uniform grid in vertical (10 m) and horizontal (10 km)



WOCE Pacific atlas http://woceatlas.ucsd.edu/index.html (Talley, 2007)

4. Oceanographic sampling: Mapping in the horizontal (DPO Section 6.4.2)



Original station data: objectively mapped to a regular grid (NOAA/ NODC)

Different types of surfaces that are commonly used:

1. Constant depth
2. Isopycnal
3. "core layer" – here salinity maximum

DPO Fig. 6.4

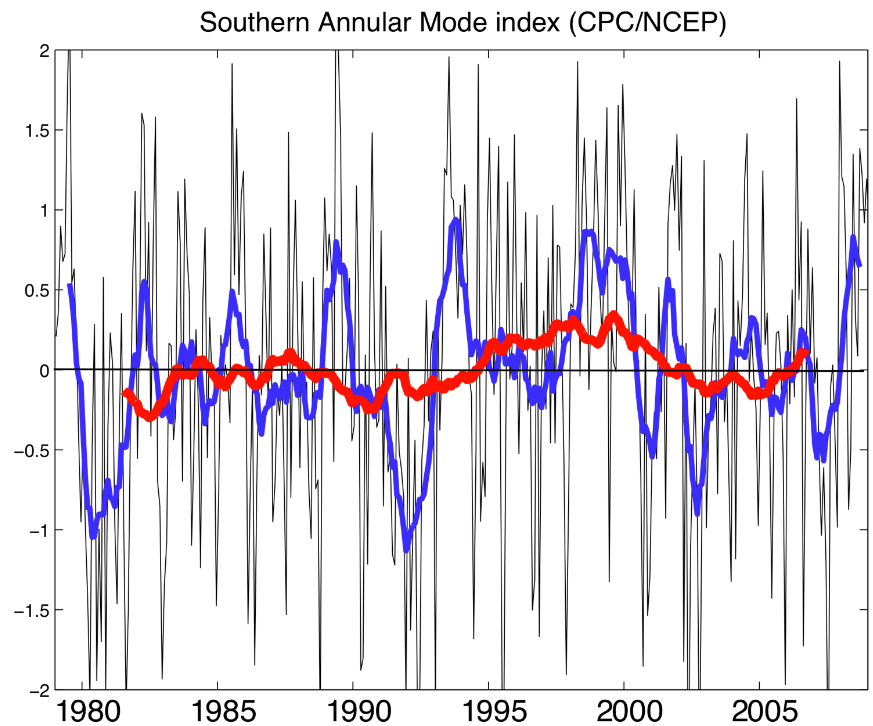
5. Filtering (DPO Section 6.5.4)

Time series or spatial sampling includes phenomena at frequencies that may not be of interest

(examples: want inertial mode but sampling includes seasonal, surface waves, etc.)

Use various filtering methods (running means with "windows" or more complicated methods of reconstructing time series) to isolate frequency of interest

Example: Time series of a climate index with 1-year and 5-year running means as filters.



DPO Fig. 6.9

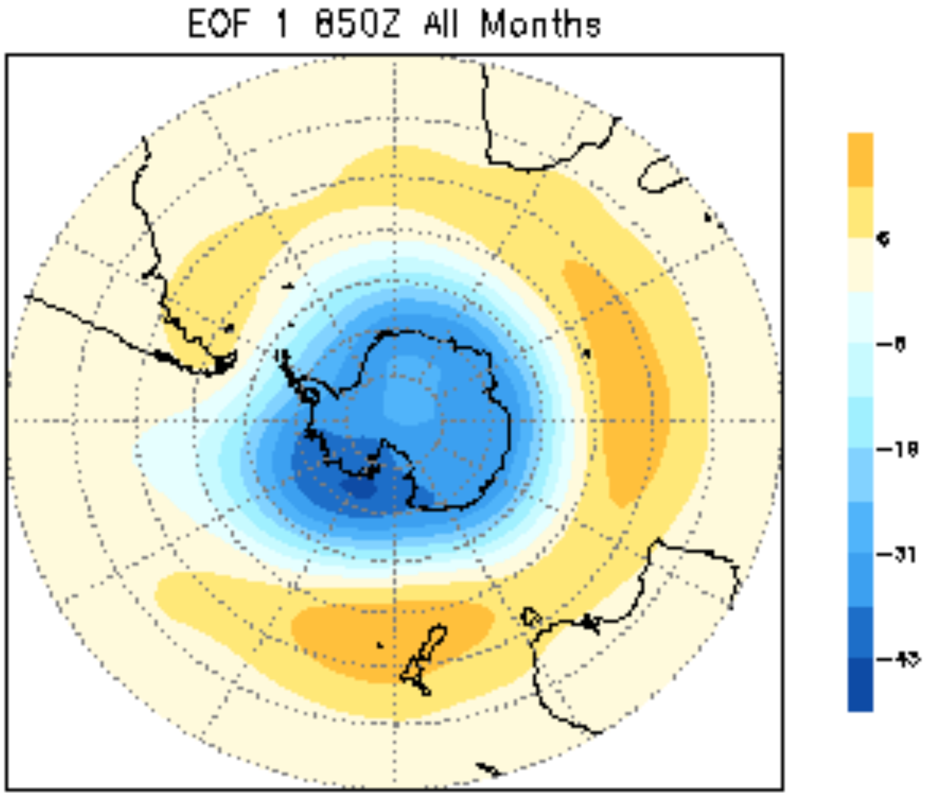
6. Space-time sampling: empirical orthogonal functions etc (DPO Section 6.6.1)

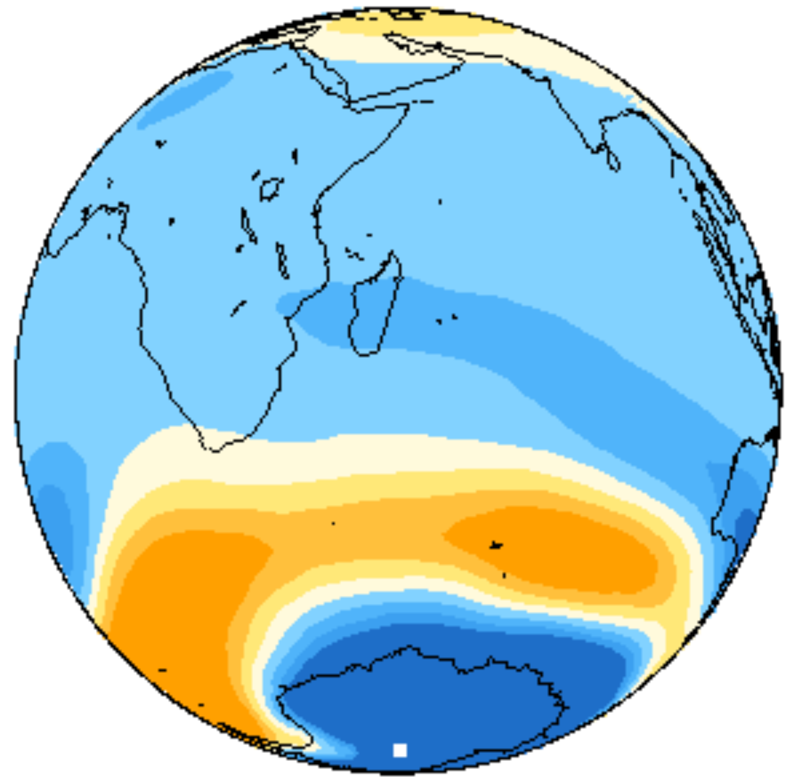
Underlying physical processes might not be best characterized by sines and cosines, especially in the spatial domain. They may be better characterized by functions that look like the basin or circulation geometries.

Empirical orthogonal function analysis: let the data decide what the basic (orthogonal) functions are that add to give the observed values.

Basic EOF analysis: obtain spatial EOFs with a time series of amplitude for each EOF. (The time series itself could be Fourier-analyzed if desired.)

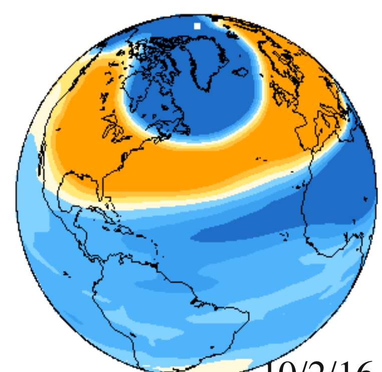
6. EOF example: Southern Annular Mode





Modes of decadal (climate) variability

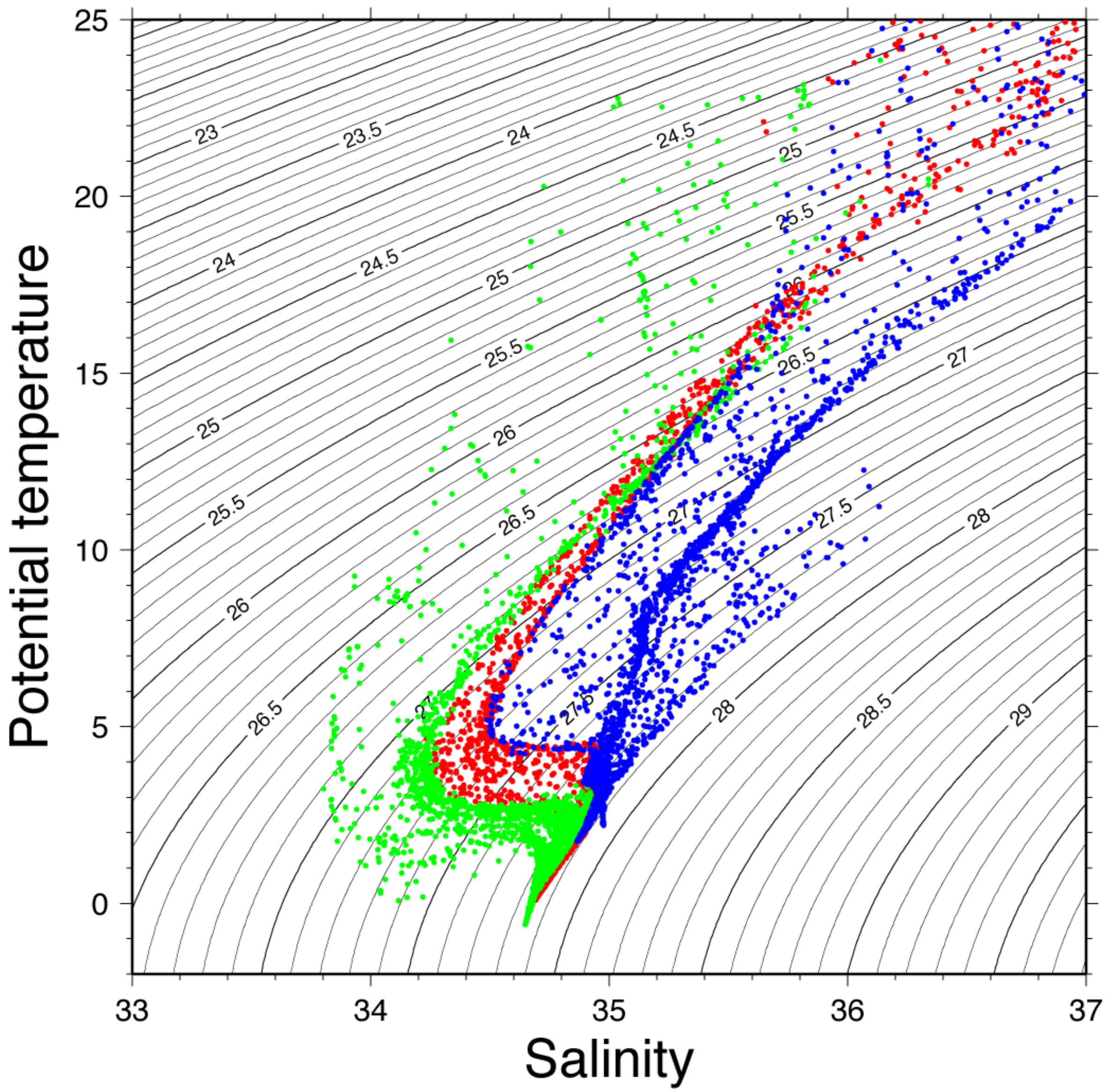
NAM



7. Water mass analysis (DPO Section 6.7.2)

- Property-property relations (e.g., theta-S)
- Volumetric property-property
- Optimum multiparameter analysis (OMP) (SIO 210 PM only)

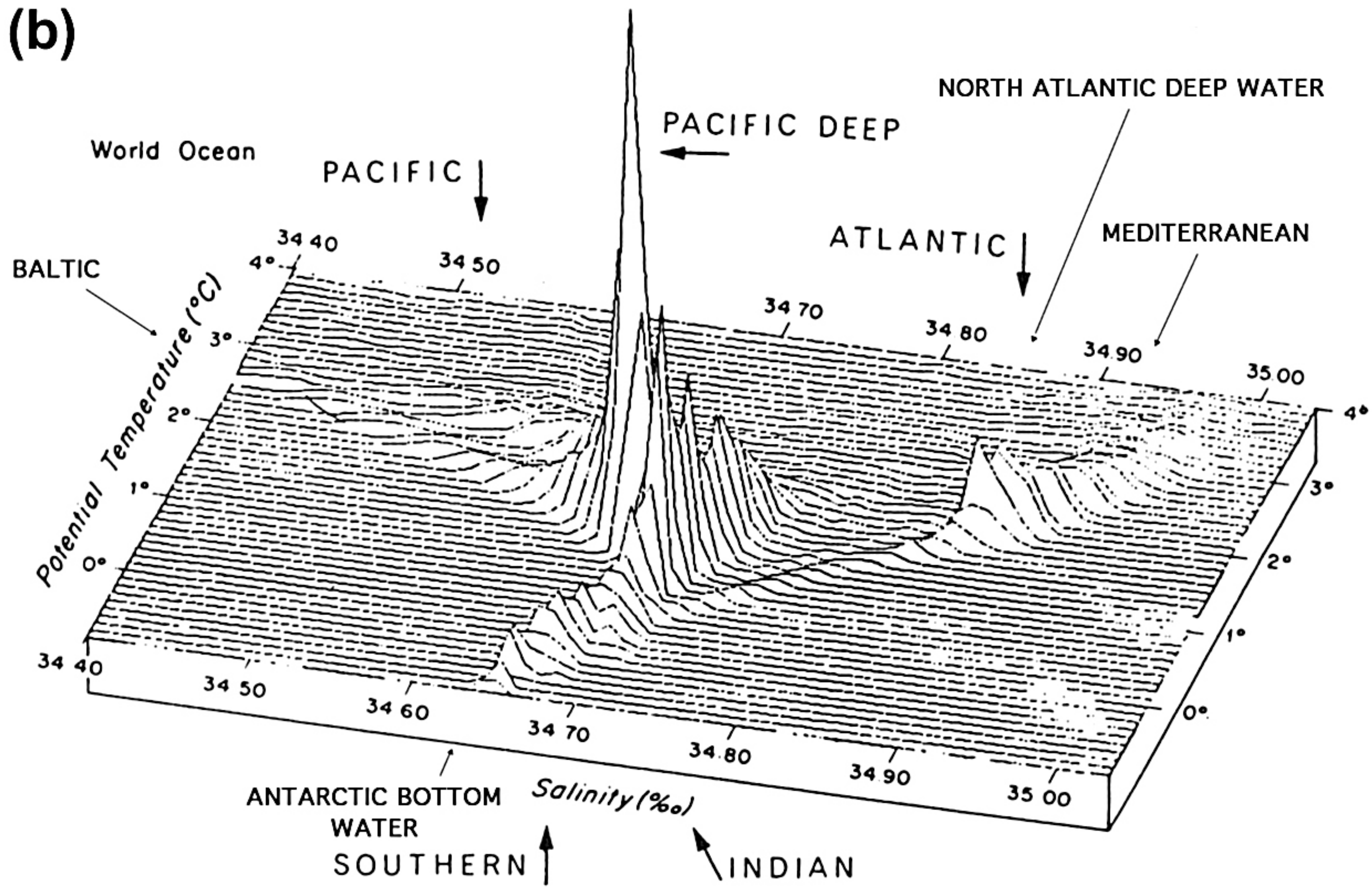
7. Water mass analysis: potential temperature-salinity diagrams (DPO Section 6.7.2)



Potential temperature-salinity diagram used in several lectures looking at Atlantic properties

DPO FIGURE 6.14

7. Water mass analysis: volumetric potential temperature-salinity (DPO Section 6.7.2)



DPO 4.17b